ECE517: Assignment 2.1

David Kirby

Due: 09 September 2020

A variation of the MMSE criterion minimizes the norm of the weight vector $\mathbf{w}$. This is a way to control the complexity of the structure. The corresponding function is

$$\mathcal{L}(\mathbf{x}, \mathbf{w}) = \mathbb{E}\left[e^2\right] + \lambda \parallel \mathbf{w} \parallel^2 \tag{1}$$

1. Make the derivation of the closed solution for $\mathbf{w}$.
2. Work out an iterative solution using the same technique as used in the Least Mean Squares algorithm.
3. Comment and compare both solutions in a short conclusion section.

The derivations must be complete and the solution should be briefly but completely explained. See the rubric for this and any other homework.

1.

$$\begin{aligned}
\mathcal{L}(\mathbf{w}) &= \mathbb{E}\left[e^2\right] + \lambda \parallel \mathbf{w} \parallel^2 && (2)\\
&= \frac{1}{N}\sum_{i=1}^{D}\left(y_n - \mathbf{w}^\top \mathbf{x}_n\right)^2 + \lambda \mathbf{w}^\top \mathbf{w}\\
&= \sum_{i=1}^{D} y_n^2 + \mathbf{w}^\top \mathbf{X}\mathbf{X}^\top \mathbf{w} - 2\mathbf{w}^\top \mathbf{X}\mathbf{y} + \lambda \mathbf{w}^\top \mathbf{w}\\
&= \sum_{i=1}^{D} y_n^2 + \mathbf{w}^\top \mathbf{R}\mathbf{w} - 2\mathbf{w}^\top \mathbf{p} + \lambda \mathbf{w}^\top \mathbf{w}\\
\nabla_{\mathbf{w}}\mathcal{L}(\mathbf{w}) &= 2\mathbf{R}\mathbf{w} - 2\mathbf{p} + 2\lambda\mathbf{w} = 0 && (3)\\
&= \mathbf{R}\mathbf{w} - \mathbf{p} + \lambda\mathbf{w}\\
&= \mathbf{R}\mathbf{w} - \mathbf{p} + \lambda\mathbf{I}\mathbf{w}\\
\mathbf{w} &= \left[\mathbf{R} + \lambda\mathbf{I}\right]^{-1}\mathbf{p} && (4)
\end{aligned}$$

2.

$$\mathbf{w}^{k+1} = \mathbf{w}^{k} - \mu\Big[\mathbf{R}\mathbf{w}^{k} - \mathbf{p} + \lambda\mathbf{I}\mathbf{w}^{k}\Big] \tag{5}$$

3. Equation (5) is an optimization that updates the parameters of equation (4) toward a maximum descent of the gradient. This equation, the Least Mean Squares Algorithm, is the least computationally burdensome, but can have issues with samples that come one at a time, or computationally complex $\mathbf{R}$ and $\mathbf{p}$.